



EUROPEAN PATENT APPLICATION
published in accordance with Art. 153(4) EPC

(43) Date of publication:
14.10.2020 Bulletin 2020/42

(51) Int Cl.:
H04R 3/00 (2006.01) H04S 7/00 (2006.01)

(21) Application number: **18895442.4**

(86) International application number:
PCT/CN2018/124244

(22) Date of filing: **27.12.2018**

(87) International publication number:
WO 2019/129127 (04.07.2019 Gazette 2019/27)

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(72) Inventors:
• **DAI, Hengjie**
Shenzhen, Guangdong 518129 (CN)
• **BAI, Hequn**
Shenzhen, Guangdong 518129 (CN)
• **XU, Dezhu**
Shenzhen, Guangdong 518129 (CN)

(30) Priority: **31.12.2017 CN 201711494923**

(74) Representative: **MERH-IP Matias Erny Reichl Hoffmann**
Patentanwälte PartG mbB
Paul-Heyse-Strasse 29
80336 München (DE)

(71) Applicant: **HUAWEI TECHNOLOGIES CO., LTD.**
Shenzhen, Guangdong 518129, (CN)

(54) **METHOD FOR MULTI-TERMINAL COOPERATIVE PLAYBACK OF AUDIO FILE AND TERMINAL**

(57) A method and a terminal for playing an audio file in a multi-terminal cooperative manner are disclosed. The method includes: obtaining, by a source terminal, an audio file, where the audio file includes an audio signal frame, and the audio signal frame includes a left channel signal and a right channel signal; obtaining, by the terminal, a central channel signal and a surround channel signal based on the left channel signal and the right channel signal; obtaining, by the terminal, a current location of a virtual sound source corresponding to the central channel signal, and generating, based on the current location and the central channel signal, a sound channel signal corresponding to the terminal in at least two sound channel signals, where the at least two sound channel signals are used to simulate a current sound field of the virtual sound source; superposing, by the terminal, the sound channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal; and playing, by the terminal, the to-be-played signal corresponding to the terminal. This method improves a sound surround effect of audio playback.

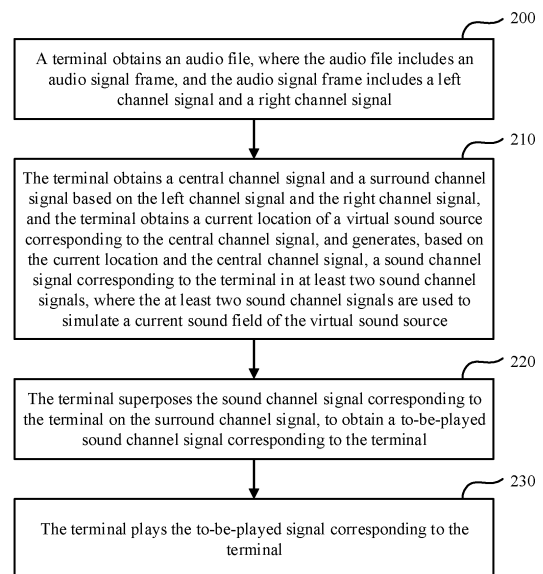


FIG. 2

Description

[0001] This application claims priority to Chinese Patent Application No. 201711494923.7, filed with the China National Intellectual Property Administration on December 31, 2017 and entitled "METHOD AND TERMINAL FOR PLAYING AUDIO FILE IN MULTI-TERMINAL CO-OPERATIVE MANNER", which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] The present invention relates to the field of terminal technologies, and in particular, to a method and a terminal for playing an audio file in a multi-terminal cooperative manner.

BACKGROUND

[0003] With rapid development of electronic technologies, terminals such as a personal computer, a smartphone, and a PDA (Personal Digital Assistant, personal digital assistant) are favored by a large quantity of users due to powerful functions of the terminals, and application of the terminals is increasingly extensive.

[0004] Currently, most terminals have an audio playback function. To ensure a playback effect of an audio file or increase playback volume of an audio file, a same audio file may be cooperatively played by using a plurality of terminals. In this case, different terminals may play different sub-channel files, to achieve an objective of improving a play effect of an audio file. The foregoing different terminals may also play the entire audio file, to achieve an effect of increasing play volume of the audio file. Usually, one terminal is selected from the plurality of terminals that perform the cooperative play operation as a source terminal, and another terminal different from the source terminal is used as a sink terminal. The source terminal sends a preset sub-channel file to each sink terminal based on pre-configured information, and after determining that transmission of the sub-channel file is completed in each terminal, controls the foregoing cooperative play process of the plurality of terminals.

[0005] However, in the prior art, because a preset sub-channel file is played in a mobile phone, a sound surround effect brought to a user is not strong.

SUMMARY

[0006] An objective of the embodiments of the present invention is to provide a method for playing an audio file in a multi-terminal cooperative manner, to improve a spatial surround effect of audio.

[0007] The foregoing objective and other objectives are achieved by using features in the independent claims. Further implementations are reflected in the dependent claims, the specification, and the accompanying drawings.

[0008] According to a first aspect, a method for playing an audio file in a multi-terminal cooperative manner is provided. The method includes:

- 5 obtaining, by a terminal, an audio file, where the audio file includes an audio signal frame, and the audio signal frame includes a left channel signal and a right channel signal;
- 10 obtaining, by the terminal, a central channel signal and a surround channel signal based on the left channel signal and the right channel signal;
- 15 obtaining, by the terminal, a current location of a virtual sound source corresponding to the central channel signal, and generating, based on the current location and the central channel signal, a sound channel signal corresponding to the terminal in at least two sound channel signals, where the at least two sound channel signals are used to simulate a current sound field of the virtual sound source;
- 20 superposing, by the terminal, the sound channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal; and
- 25 playing, by the terminal, the to-be-played signal corresponding to the terminal.

[0009] The foregoing method may be performed by a source terminal, or may be executed by a sink terminal.

[0010] The signal may be understood as audio data, for example, to-be-processed audio data. For example, the sound channel signal may be understood as sound channel audio data, and the signal frame may be understood as a data frame.

[0011] The sound channel signal corresponding to the terminal means that there are at least two terminals in a cooperative playing system, and the terminals play different channel signals. A correspondence between the terminal and the channel signal may be implemented by using a preset correspondence, for example, a correspondence between a sequence number of the terminal and a sequence number of a sound channel. Alternatively, the sound channel signal corresponding to the terminal may be determined based on a relative location relationship between the terminal and another terminal in the at least two terminals.

[0012] The simulating a current sound field of the virtual sound source may mean that simulating a sound field that is generated at a human ear location when the virtual sound source is at the current location. The human ear location may be detected by the source terminal, or may be preset.

[0013] With reference to the first aspect, in a first possible implementation of the first aspect, the terminal is a source terminal, and the method further includes:

- 50 controlling, by the source terminal, at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played signal corresponding to the source terminal in the at least two to-be-played

sound channel signals, to control the at least one sink terminal to cooperatively play the at least two to-be-played sound channel signals with the terminal.

[0014] The at least one sink terminal may be at least two sink terminals, at least three sink terminals, or at least four sink terminals.

[0015] The at least one sink terminal is in one-to-one correspondence with the at least one to-be-played sound channel signal, that is, one terminal in the at least one sink terminal corresponds to one sound channel signal in the at least one to-be-played sound channel signal. The controlling at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played signal corresponding to the source terminal in the at least two to-be-played sound channel signals may specifically include: controlling the at least one sink terminal to play a respective sound channel signal corresponding to the at least one sink terminal in the at least one to-be-played sound channel signal.

[0016] With reference to the first aspect or the first possible implementation of the first aspect, in a second possible implementation of the first aspect, the obtaining a current location of a virtual sound source corresponding to the central channel signal includes:

obtaining a movement speed of the virtual sound source and moment information of the audio signal frame; and

determining, based on a preset movement track of the virtual sound source, the movement speed, and the moment information, the current location of the virtual sound source on the movement track.

[0017] The moment information may be determined based on a frame sequence number of the audio signal frame.

[0018] The determining a current location of a virtual sound source may include: determining the current location based on a difference between the moment information and stored previous moment information before the moment information, a location corresponding to the stored previous time information on the movement track, and the movement speed. The method may further include: storing the current location and the moment information, where the current location corresponds to the moment information.

[0019] With reference to the second possible implementation of the first aspect, in a third possible implementation of the first aspect, the audio signal frame includes music data, and the obtaining a movement speed of the virtual sound source includes:

determining rhythm information of music indicated by the audio signal frame; and

determining the movement speed based on the rhythm information, where a faster rhythm indicated by the rhythm information indicates a faster movement speed.

[0020] The music indicated by the audio signal frame is music generated by playing the audio signal frame.

[0021] With reference to the third possible implementation of the first aspect, in a fourth possible implementation of the first aspect, the determining rhythm information of music indicated by the audio signal frame includes: determining the rhyme information based on the audio signal frame and N signal frames before the audio signal frame in the audio file, where N is an integer greater than 0.

[0022] With reference to the second possible implementation, the third possible implementation, or the fourth possible implementation of the first aspect, in a fifth possible implementation of the first aspect, the movement track is a circle that rotates around a preset human ear location.

[0023] With reference to the fifth possible implementation of the first aspect, in a sixth possible implementation of the first aspect, the terminal is the source terminal, and the source terminal or the at least one sink terminal controlled by the source terminal is located in a plane in which the circle is located. According to a second aspect, a terminal for playing an audio file in a multi-terminal cooperative manner is provided. The terminal includes:

a first obtaining unit, configured to obtain an audio file, where the audio file includes an audio signal frame, and the audio signal frame includes a left channel signal and a right channel signal;

a second obtaining unit, configured to obtain a central channel signal and a surround channel signal based on the left channel signal and the right channel signal;

a generation unit, configured to generate a current location of a virtual sound source corresponding to the central channel signal, and generate, based on the current location and the central channel signal, a sound channel signal corresponding to the terminal in at least two sound channel signals, where the at least two sound channel signals are used to simulate a current sound field of the virtual sound source;

a superposition unit, configured to superpose the sound channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal; and

a playback unit, configured to play the to-be-played signal corresponding to the terminal.

[0024] With reference to the second aspect, in a first possible implementation of the second aspect, the terminal is a source terminal, and the source terminal further includes:

a controlling unit, configured to control at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played signal corresponding to the source terminal in the at least two to-be-played sound channel signals, to control the at least one sink

terminal to cooperatively play the at least two to-be-played sound channel signals with the terminal.

[0025] With reference to the second aspect or the first possible implementation of the second aspect, in a second possible implementation of the second aspect, the generation unit is configured to:

obtain a movement speed of the virtual sound source and moment information of the audio signal frame; and
determine, based on a preset movement track of the virtual sound source, the movement speed, and the moment information, the current location of the virtual sound source on the movement track.

[0026] With reference to the second possible implementation of the second aspect, in a third possible implementation of the second aspect, the audio signal frame includes music data, and the generation unit is configured to:

determine rhythm information of music indicated by the audio signal frame; and
determine the movement speed based on the rhythm information, where a faster rhythm indicated by the rhythm information indicates a faster movement speed.

[0027] With reference to the third possible implementation of the second aspect, in a fourth possible implementation of the second aspect, the generation unit is configured to:

determine the rhyme information based on the audio signal frame and N signal frames before the audio signal frame in the audio file, where N is an integer greater than 0.

[0028] With reference to the second possible implementation, the third possible implementation, or the fourth possible implementation of the second aspect, in a fifth possible implementation of the second aspect, the movement track is a circle that rotates around a preset human ear location.

[0029] With reference to the fifth possible implementation of the second aspect, in a sixth possible implementation of the second aspect, the terminal is the source terminal, and the source terminal or the at least one sink terminal controlled by the source terminal is located in a plane in which the circle is located.

[0030] According to a third aspect, a terminal for playing an audio file in a multi-terminal cooperative manner is provided. The terminal includes a memory and a processor, where
the memory is configured to store a set of executable code; and
the processor is configured to execute the executable code stored in the memory, to perform any one of the first aspect or the possible implementations of the first aspect.

[0031] According to a fourth aspect, a storage medium is provided. The storage medium stores executable code, and when the executable code is executed, any one of the first aspect or the possible implementations of the first aspect may be performed.

[0032] According to a fifth aspect, a computer program is provided. The computer program may perform any one of the first aspect or the possible implementations of the first aspect.

[0033] According to a sixth aspect, a computer program product is provided. The computer program product includes an instruction that can execute any one of the first aspect or the possible implementations of the first aspect.

BRIEF DESCRIPTION OF DRAWINGS

[0034] To describe the technical solutions in the embodiments of the present invention more clearly, the following briefly describes the accompanying drawings required for describing the embodiments.

FIG. 1 is an architectural diagram of a system for playing an audio file in a multi-terminal cooperative manner according to an embodiment of the present invention;

FIG. 2 is a flowchart of a method for playing an audio file in a multi-terminal cooperative manner according to an embodiment of the present invention;

FIG. 3 is a schematic structural diagram of a terminal configured to play an audio file in a multi-terminal cooperative manner according to an embodiment of the present invention; and

FIG. 4 is a schematic structural diagram of a terminal configured to play an audio file in a multi-terminal cooperative manner according to an embodiment of the present invention.

DESCRIPTION OF EMBODIMENTS

[0035] The following clearly and completely describes the technical solutions in the embodiments of the present invention with reference to the accompanying drawings in the embodiments of the present invention.

[0036] FIG. 1 is an architectural diagram of a system according to an embodiment of the present invention. A source terminal may cooperatively play an audio file with one sink terminal, or may cooperatively play an audio file with a plurality of sink terminals. It should be noted that in this embodiment of the present invention, a plurality of terminals may be at least two terminals, at least three terminals, at least four terminals, three terminals, four terminals, five terminals, six terminals, seven terminals, or eight terminals.

[0037] In this embodiment of the present invention, terminals participating in cooperative playing of an audio file establish a connection to each other in a wired or wireless manner. A person skilled in the art may under-

stand that the "terminal" and a "terminal device" used herein include a device that has a wireless signal receiver having no transmit capability, and further include a device that has receiving and transmitting hardware having a capability of performing bidirectional communication on a bidirectional communication link. Such a device may include: a cellular device or another communication device that has a single line display or a multiline display or has no multiline display; a PCS (Personal Communications Service, personal communications service), where voice, data processing, fax, and/or data communication capabilities may be combined in the PCS; a PDA (Personal Digital Assistant, personal digital assistant) that may include a radio frequency receiver, a pager, Internet/intranet access, a web browser, a notepad, a calendar, and/or a GPS (Global Positioning System, global positioning system) receiver; and a conventional laptop and/or palmtop computer or another device that has and/or includes a radio frequency receiver. The "terminal" and the "terminal device" used herein may be portable, transportable, and installed in a vehicle (a vehicle on air, sea, and/or land), or suitable for running and/or configured to run locally, and/or run at any other location on Earth and/or in space in a distributed form. The "terminal" and the "terminal device" used herein may further be a communications terminal, an Internet access terminal, and a music/video playing terminal, for example, may be a PDA, a MID (Mobile Internet Device, mobile Internet device), and/or a mobile phone having a music/video playing function, or may be a device such as a smart television or a set-top box. After connections between the terminals participating in cooperative playing of the audio file are established, the terminals need to be configured, to be specific, a source terminal and a sink terminal are configured, where the source terminal and the sink terminal are the terminals. The source terminal may be specified by a user, or may be determined based on a pre-configuration. Usually, any terminal in terminals including a specified audio file is used as a source terminal, and another terminal participating in cooperative playing of the audio file that is different from the source terminal is used as a sink terminal.

[0038] After the source terminal and the sink terminal are configured, the source terminal serves as a playback control unit to transmit a multi-channel audio file (the audio file includes a channel signal) and deliver a control instruction to the sink terminal. In this embodiment of the present invention, the user may send a control instruction by using the source terminal to another terminal in a terminal group, where the control instruction includes an instruction such as a playback instruction or a playback stop instruction. The source terminal and the sink terminal may perform one or more types of cooperative sound effect processing based on a song and a playing mode that are selected by the user. There may be one or more sink terminals participating in cooperative playing of the audio file.

[0039] Referring to FIG. 2, in an embodiment of the

present invention, an execution body may be a source terminal, a sink terminal, or a non-terminal-type computer device. The following uses the source terminal as an example for description. A process in which a plurality of terminals cooperatively play an audio file is as follows:

[0040] Step 200: A terminal obtains the audio file, where the audio file includes an audio signal frame, and the audio signal frame includes a left channel signal and a right channel signal.

[0041] The signal may be understood as audio data, for example, to-be-processed audio data. For example, the sound channel signal may be understood as sound channel audio data, and the signal frame may be understood as a data frame.

[0042] Step 210: The terminal obtains a central channel signal and a surround channel signal based on the left channel signal and the right channel signal, and the terminal obtains a current location of a virtual sound source corresponding to the central channel signal, and generates, based on the current location and the central channel signal, a sound channel signal corresponding to the terminal in at least two sound channel signals, where the at least two sound channel signals are used to simulate a current sound field of the virtual sound source.

[0043] The sound channel signal corresponding to the terminal may be generated by using a speaker virtual mapping technology. This technology encodes the virtual sound source to an Ambisonic domain through spherical harmonic decomposition based on a location of the virtual sound source in a Cartesian coordinate system, calculates a decoding matrix based on a location of a playback speaker, and decodes the encoded file to the speaker for playback.

[0044] During specific implementation, the generating the at least two sound channel signals based on the current location and the central channel signal may include: generating the at least two sound channel signals based on the current location, the central channel signal, a human ear location, and location distribution of the terminal group. During specific implementation, a source terminal may control each terminal in the terminal group to send an ultrasonic wave, and each terminal calculates a distance between terminals based on the ultrasonic wave, to obtain location distribution of the terminal group. The terminal group includes the source terminal and at least one sink terminal. For example, a source terminal A instructs a terminal B to send an ultrasonic wave, and after sending the ultrasonic wave, the terminal B sends, to the source terminal A, a time at which the ultrasonic wave is sent. The source terminal A calculates a distance between the terminal B and the terminal A based on the time at which the terminal B sends the ultrasonic wave and a time at which the terminal A receives the ultrasonic wave. In this way, location distribution of the terminals in the terminal group is obtained. In another implementation, location distribution of the terminal group is preset. Similarly, when the terminal group is used to play audio, a user may be required to place the terminal group based

on a preset location.

[0045] The obtaining a current location of a virtual sound source corresponding to the central channel signal may include: obtaining a movement speed of the virtual sound source and moment information of the audio signal frame; and determining, based on a preset movement track of the virtual sound source, the movement speed, and the moment information, the current location of the virtual sound source on the movement track.

[0046] In a possible implementation, the audio signal frame includes music data, and the obtaining a movement speed of the virtual sound source may include: determining rhythm information of music indicated by the audio signal frame; and determining the movement speed based on the rhythm information, where a faster rhythm indicated by the rhythm information indicates a faster movement speed. The determining rhythm information of music indicated by the audio signal frame may include: determining the rhyme information based on the audio signal frame and N signal frames before the audio signal frame in the audio file, where N is an integer greater than 0.

[0047] In a possible implementation, the movement track may be a circle that rotates around a human ear location. Further, the source terminal or the at least one sink terminal controlled by the source terminal may be located in a plane in which the circle is located. Alternatively, the source terminal and the at least one sink terminal may be located in a plane in which the circle is located. Certainly, specifically, the source terminal or the at least one sink terminal may be located at the circle. During actual application, the human ear location may be a location at which the user performs entering by using a UI interface of the source terminal. Alternatively, the human ear location may be a preset relative location relative to the source terminal and/or a specific sink terminal (or some sink terminals).

[0048] Alternatively, a terminal (a source terminal or a sink terminal) photographs a head picture of a user, to determine a listening location of the user as the human ear location.

[0049] Step 220: The terminal superposes a channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal.

[0050] Step 203: The terminal plays the to-be-played signal corresponding to the terminal.

[0051] When the terminal is a source terminal, the method may further include: controlling, by the source terminal, at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played signal corresponding to the source terminal in the at least two to-be-played sound channel signals, to control the at least one sink terminal to cooperatively play the at least two to-be-played sound channel signals with the terminal. It may be understood that for more content, refer to related descriptions in content of the present invention. Details are not described herein again.

[0052] An embodiment of the present invention further provides a system for playing an audio file in a multi-terminal cooperative manner. The system includes the source terminal that performs the foregoing method that may be performed by the source terminal, and the sink terminal that performs the foregoing method that may be performed by the sink terminal. It should be noted that if it is not specially noted that a method is performed by the source terminal, the method may be performed by the source terminal, or may be performed by the sink terminal.

[0053] The following provides a description with reference to a specific application scenario. The application scenario may be as follows: When a plurality of people gather, a plurality of mobile phones are placed at a predetermined location around a gathering site, and are simultaneously connected to a same Wi-Fi hotspot. The mobile phones use the Wi-Fi hotspot to communicate with each other, play music, and make a human voice (a central channel signal) act as a rhythmic movement element between devices. When a user chooses to play relatively comfortable music, the movement element moves slowly between devices, bringing an elegant party experience. When the user chooses to play a song with a strong rhythm, the movement element has a quick rhythm based on the rhythm of the song, thereby increasing a sense of rhythm for the party.

[0054] Herein, an example in which the system for playing an audio file in a multi-terminal cooperative manner includes three terminals (a terminal A, a terminal B, and a terminal C) and the terminal A, the terminal B, and the terminal C cooperatively play an audio file is used to describe a method procedure for cooperatively playing an audio file by a plurality of terminals and a system for playing an audio file in a multi-terminal cooperative manner. The procedure includes:

Step 0: Establish a connection relationship between the terminal A, the terminal B, and the terminal C, where the terminal A is configured as a source terminal, and the terminal B and the terminal C are configured as sink terminals.

Step 1: The terminal A obtains an audio file, and divides the audio file into signal frames of a same size.

[0055] That sizes are the same may mean that quantities of sampling points in all frames are the same. The audio file may be a stereo audio file, a 5.1-channel audio file, a 7.1-channel audio file, or the like, and these audio files are not further listed one by one herein.

[0056] Step 2: The terminal A obtains a user-preset movement curve and an initial location of a virtual sound source on the movement curve, where the movement curve may be a circle, and the terminal A, the terminal B, and the terminal C are located in a plane in which the circle is located. The reason is that simulation of a sound field in the plane is easier than that in space.

[0057] The movement curve may be a function about a time and three-dimensional coordinates. The movement curve is a movement curve of the virtual sound source.

[0058] A central extraction technology is to extract the virtual central channel signal from a dual-channel input sound source in a channel upmixing manner. There are different methods for implementing channel upmixing. Some methods use matrix decoding that is performed in time domain. Some methods are based on signal correlation. For example, it is assumed that left, central, and right signals (L, C, and R) obtained after the left and right channel signals are upmixed are not correlated. In this case, the central channel signal is extracted in frequency domain.

[0059] Extraction of the surround channel signal may be extracting anticorrelated surrounding information in time domain by using a left and right channel de-correlation method. For example, an azimuth is calculated based on energy of left and right channels, and weighting factors of the left and right channels are calculated based on azimuth information, for example, $SL = a * L + b * R$, where a and b are calculated weighting factors. Specific implementation may be a surround sound $S = L * 0.4 - R * 0.3$. Step 3: In a process in which the virtual sound source moves, the terminal A detects rhythm information of music indicated by a current audio signal frame, and updates a movement speed based on the rhythm information. The faster the rhythm information is, the faster the movement speed is.

[0060] It should be noted that if the rhyme information is detected for the first time, it means that the movement speed is not updated previously. In this case, for the rhyme information detected for the first time, the movement speed is determined based on the detected rhyme information.

[0061] Specifically, a manner in which movement information is updated may be determining, based on the rhythm information, a movement speed corresponding to the rhythm information, where the movement speed is used to update the movement speed. Alternatively, after the movement speed corresponding to the rhyme information is determined, a weight sum of the movement speed and a movement speed that corresponds to previous rhyme information and that is determined based on the rhyme information last time may be used as the movement speed for update. In this case, in step 2, an initial value of the movement speed needs to be obtained.

[0062] Rhythm information of music indicated by a current audio signal frame and N frames before the current audio signal frame may be detected and used as rhythm information of music indicated by the current audio signal frame, where N may be 10.

[0063] Step 4: The terminal A determines the current location of the virtual sound source based on the moment information indicated by a sequence number of the current audio signal frame, the moment information corresponding to a previous audio signal frame, a location of

a previous virtual sound source, and an updated movement speed. The current location may be represented by using a three-dimensional coordinate value. The location of the virtual sound source may be understood as a location of the human sound or the instrument sound.

[0064] The moment information corresponding to the previous audio signal frame and the location of the previous virtual sound source may be, when the moment speed is updated last time, moment information corresponding to an analyzed audio signal frame and a location of a determined virtual sound source.

[0065] Specifically, the terminal A may obtain a difference between the moment information indicated by the sequence number of the current audio signal frame and the moment information corresponding to the previous audio signal frame, and then determine the current location, where a displacement of the current location relative to the previous location along the movement track is a product of the difference and the updated movement speed.

[0066] Step 5: The terminal A obtains the central channel signal and the surround channel signal based on the current audio signal frame in the audio file.

[0067] Step 6: The terminal A processes the central channel signal based on the current location of the virtual sound source, to obtain a channel signal corresponding to the terminal A in three channel signals. The three sub-channel signals are used to simulate a sound field that is at a human ear location when the virtual sound source is at the current location.

[0068] Step 7: The terminal A superposes the channel signal corresponding to the terminal A on the surround channel signal, to obtain a to-be-played sound channel signal used for playing by the terminal A.

[0069] Step 8: Similar to that the terminal A obtains the to-be-played sound channel signal used for playing by the terminal A, the terminal B obtains a to-be-played sound channel signal used for playing by the terminal B, and the terminal C obtains a to-be-played signal used for playing by the terminal C.

[0070] Step 9: The terminal A controls the terminal A to play the to-be-played sound channel signal used for playing by the terminal A, controls the terminal B to play the to-be-played sound channel signal used for playing by the terminal B, and controls the terminal C to play the to-be-played sound channel signal used for playing by the terminal C.

[0071] Step 10: When all signal frames in the audio file are processed, the procedure ends; otherwise, step 3 is performed.

[0072] As shown in FIG. 3, an embodiment of the present invention provides a terminal 300 for playing an audio file in a multi-terminal cooperative manner. The terminal 300 is a source terminal, and the terminal 300 may include a first obtaining unit 301, a second obtaining unit 302, a generation unit 303, a superposition unit 304, and a sending unit 305. Operations performed by the units in the apparatus may be implemented by using soft-

ware, and may be used as a software module located in a memory of the terminal 300 and invoked and executed by a processor. The operations performed by the units in the apparatus may be alternatively implemented by using a hardware chip.

[0073] The first obtaining unit 301 is configured to obtain an audio file, where the audio file includes an audio signal frame, and the audio signal frame includes a left channel signal and a right channel signal.

[0074] The second obtaining unit 302 is configured to obtain a central channel signal and a surround channel signal based on the left channel signal and the right channel signal.

[0075] The generation unit 303 is configured to generate a current location of a virtual sound source corresponding to the central channel signal, and generate, based on the current location and the central channel signal, a sound channel signal corresponding to terminal in at least two sound channel signals, where the at least two sound channel signals are used to simulate a current sound field of the virtual sound source.

[0076] The generation unit 303 may be configured to obtain a movement speed of the virtual sound source and moment information of the audio signal frame, and determine, based on a preset movement track of the virtual sound source, the movement speed, and the moment information, the current location of the virtual sound source on the movement track.

[0077] In a possible implementation, the audio signal frame includes music data, and the generation unit 303 may be configured to: determine rhythm information of music indicated by the audio signal frame; and determine the movement speed based on the rhythm information, where a faster rhythm indicated by the rhythm information indicates a faster movement speed. The generation unit 303 may be configured to: determine the rhyme information based on the audio signal frame and N signal frames before the audio signal frame in the audio file, where N is an integer greater than 0. The movement track may be a circle that rotates around a preset human ear location. The source terminal or the at least one sink terminal may be located in a plane in which the circle is located.

[0078] In a possible implementation, the generation unit 303 generates the at least two sound channel signals based on the current location and the central channel signal only when the current location does not overlap a location of a playback terminal, where the playback terminal is the source terminal, or the playback terminal is one of the at least one sink terminal.

[0079] The superposition unit 304 is configured to superpose the sound channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal.

[0080] The playback unit 305 is configured to play the to-be-played signal corresponding to the terminal.

[0081] When the terminal is the source terminal, the terminal may further include: a controlling unit, configured

to control at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played sound signal corresponding to the source terminal in the at least two to-be-played sound channel signals, to control the at least one sink terminal to cooperatively play the at least two to-be-played sound channel signals with the terminal.

[0082] It may be understood that, for more operations performed by the units of the terminal in this embodiment, refer to related descriptions in the foregoing method embodiments and the summary. Details are not described herein again.

[0083] FIG. 4 is a schematic structural diagram of a terminal 400 configured to play an audio file in a multi-terminal cooperative manner according to an embodiment of the present invention. As shown in FIG. 4, the terminal 400 may be used as an implementation of the terminal 300. The terminal 400 includes a processor 402, a memory 404, an input/output interface 406, a communications interface 408, and a bus 410. The processor 402, the memory 404, the input/output interface 406, and the communications interface 408 implement a mutual communication connection by using the bus 410.

[0084] The processor 402 may be a general-purpose central processing unit (Central Processing Unit, CPU), a microprocessor, an application-specific integrated circuit (Application Specific Integrated Circuit, ASIC), or one or more integrated circuits, and is configured to execute a related program, to implement functions that units included in the device 300 provided in the embodiments of the present invention needs to perform, or perform the photographing methods provided in the method embodiments and the summary of the present invention. The processor 402 may be an integrated circuit chip and has a signal processing capability. In an implementation process, steps in the foregoing methods can be implemented by using a hardware integrated logical circuit in the processor 402, or by using an instruction in a form of software. The processor 402 may be a general purpose processor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), a field programmable gate array (FPGA) or another programmable logic device, a discrete gate or a transistor logic device, or a discrete hardware component. The processor 402 may implement or perform the methods, the steps, and logical block diagrams that are disclosed in the embodiments of the present invention. The general purpose processor may be a microprocessor, or the processor may be any conventional processor or the like. The steps of the methods disclosed with reference to the embodiments of the present invention may be directly performed by a hardware decoding processor, or may be performed by using a combination of hardware and software units in the decoding processor. A software unit may be located in a mature storage medium in the art, such as a random access memory, a flash memory, a read-only memory, a programmable read-only memory, an electrically erasable programmable memory, or a register. The storage medium is located

in the memory 404, and the processor 402 reads information in the memory 404 and completes the steps in the foregoing methods in combination with hardware of the processor 402.

[0085] The memory 404 may be a read-only memory (Read Only Memory, ROM), a static storage device, a dynamic storage device, or a random access memory (Random Access Memory, RAM). The memory 404 may store an operating system and another application program. When functions that need to be performed by the units included in the terminal 300 provided in the embodiments of the present invention are implemented by using software or firmware, or when the photographing methods provided in the method embodiments and the summary of the present invention are performed, program code used to implement the technical solutions provided in the embodiments of the present invention is stored in the memory 404, and the processor 402 performs the operations that need to be performed by the units included in the terminal 300, or performs the photographing methods provided in the method embodiments of the present invention.

[0086] The input/output interface 406 is configured to receive input data and information, and output data such as an operation result.

[0087] The communications interface 408 uses a transceiver apparatus, for example but not limited to, a transceiver, to implement communication between the terminal 400 and another device or a communications network.

[0088] The bus 410 may include a path, for transmitting information between the components (for example, the processor 402, the memory 404, the input/output interface 406, and the communications interface 408) of the terminal 400.

[0089] It should be noted that although the terminal 400 shown in FIG. 4 shows only the processor 402, the memory 404, the input/output interface 406, the communications interface 408, and the bus 410, in a specific implementation process, a person skilled in the art should understand that the terminal 400 further includes other components required for implementing normal running, for example, a display, a camera, and a gyroscope sensor. In addition, according to a specific requirement, a person skilled in the art should understand that the terminal 400 may further include a hardware component for implementing another additional function. In addition, a person skilled in the art should understand that the terminal 400 may include only a component essential for implementing the embodiments of the present invention, but not necessarily include all the components shown in FIG. 4.

[0090] It may be understood that, for operations performed by the elements of the terminal in this embodiment, refer to the related descriptions in the foregoing method embodiments and the summary. Details are not described herein again.

[0091] It should be noted that, for simplicity of descrip-

tion, the foregoing method embodiments are expressed as a combination of a series of actions. However, a person skilled in the art should appreciate that the present invention is not limited to the described action sequence.

That is because according to the present invention, some steps may be performed in another sequence or performed simultaneously. In addition, a person skilled in the art should also appreciate that, actions and units in the specification are not necessary for the present invention.

[0092] A person of ordinary skill in the art may understand that all or some of the procedures of the methods in the foregoing embodiments may be implemented by a computer program instructing relevant hardware. The program may be stored in a computer-readable storage medium. When the program is run, the procedures of the methods in the embodiments are performed. The foregoing storage medium may be a magnetic disk, an optical disc, a read-only memory (ROM: Read-Only Memory), a random access memory (RAM: Random Access Memory), or the like.

[0093] The present invention is described with reference to the flowcharts and/or block diagrams of the method, the device (system), and the computer program product according to the embodiments of the present invention. It should be understood that computer program instructions may be used to implement each procedure and/or each block in the flowcharts and/or the block diagrams and a combination of a procedure and/or a block in the flowcharts and/or the block diagrams. These computer program instructions may be provided for a processor of a general-purpose computer, a special-purpose computer, an embedded processor, or another programmable data processing device to generate a machine, so that the instructions executed by the processor of the computer or the another programmable data processing device generate an apparatus for implementing a specific function in one or more procedures in the flowcharts and/or in one or more blocks in the block diagrams.

[0094] These computer program instructions may alternatively be stored in a computer readable memory that can instruct a computer or another programmable data processing device to work in a specific manner, so that the instructions stored in the computer readable memory generate an artifact that includes an instruction apparatus. The instruction apparatus implements a specific function in one or more procedures in the flowcharts and/or in one or more blocks in the block diagrams.

[0095] These computer program instructions may alternatively be loaded onto a computer or another programmable data processing device, so that a series of operations and steps are performed on the computer or the another programmable device, thereby generating computer-implemented processing. Therefore, the instructions executed on the computer or the another programmable device provide steps for implementing a specific function in one or more procedures in the flowcharts and/or in one or more blocks in the block diagrams.

[0096] Although some preferred embodiments of the present invention have been described, a person skilled in the art can make additional changes and modifications to these embodiments once they learn of the basic inventive concept. Therefore, the following claims are intended to be construed as to cover the preferred embodiments and all changes and modifications falling within the scope of the present invention.

[0097] Obviously, a person skilled in the art can make various modifications and variations to the embodiments of the present invention without departing from the spirit and scope of the embodiments of the present invention. In this way, the present invention is intended to cover these modifications and variations provided that these modifications and variations of the embodiments of the present invention fall within the scope of the claims and equivalent technologies of the claims of the present invention.

Claims

1. A method for playing an audio file in a multi-terminal cooperative manner, comprising:

obtaining, by a terminal, an audio file, wherein the audio file comprises an audio signal frame, and the audio signal frame comprises a left channel signal and a right channel signal;
 obtaining, by the terminal, a central channel signal and a surround channel signal based on the left channel signal and the right channel signal;
 obtaining, by the terminal, a current location of a virtual sound source corresponding to the central channel signal, and generating, based on the current location and the central channel signal, a sound channel signal corresponding to the terminal in at least two sound channel signals, wherein the at least two sound channel signals are used to simulate a current sound field of the virtual sound source;
 superposing, by the terminal, the sound channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal; and
 playing, by the terminal, the to-be-played signal corresponding to the terminal.

2. The method according to claim 1, wherein the terminal is a source terminal, and the method further comprises:

controlling, by the source terminal, at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played signal corresponding to the source terminal in the at least two to-be-played sound channel signals, to control the at least one sink terminal to cooperatively play

the at least two to-be-played sound channel signals with the terminal.

3. The method according to claim 1 or 2, wherein the obtaining a current location of a virtual sound source corresponding to the central channel signal comprises:

obtaining a movement speed of the virtual sound source and moment information of the audio signal frame; and
 determining, based on a preset movement track of the virtual sound source, the movement speed, and the moment information, the current location of the virtual sound source on the movement track.

4. The method according to claim 3, wherein the audio signal frame comprises music data, and the obtaining a movement speed of the virtual sound source comprises:

determining rhythm information of music indicated by the audio signal frame; and
 determining the movement speed based on the rhythm information, wherein a faster rhythm indicated by the rhythm information indicates a faster movement speed.

5. The method according to claim 4, wherein the determining rhythm information of music indicated by the audio signal frame comprises:

determining the rhyme information based on the audio signal frame and N signal frames before the audio signal frame in the audio file, wherein N is an integer greater than 0.

6. The method according to any one of claims 3 to 5, wherein the movement track is a circle that rotates around a preset human ear location.

7. The method according to claim 6, wherein the terminal is the source terminal, and the source terminal or the at least one sink terminal controlled by the source terminal is located in a plane in which the circle is located.

8. A terminal for playing an audio file in a multi-terminal cooperative manner, wherein the terminal comprises:

a first obtaining unit, configured to obtain an audio file, wherein the audio file comprises an audio signal frame, and the audio signal frame comprises a left channel signal and a right channel signal;

a second obtaining unit, configured to obtain a central channel signal and a surround channel

- signal based on the left channel signal and the right channel signal;
 a generation unit, configured to generate a current location of a virtual sound source corresponding to the central channel signal, and generate, based on the current location and the central channel signal, a sound channel signal corresponding to the terminal in at least two sound channel signals, wherein the at least two sound channel signals are used to simulate a current sound field of the virtual sound source;
 a superposition unit, configured to superpose the sound channel signal corresponding to the terminal on the surround channel signal, to obtain a to-be-played sound channel signal corresponding to the terminal; and
 a playback unit, configured to play the to-be-played signal corresponding to the terminal.
9. The terminal according to claim 8, wherein the terminal is a source terminal, and the source terminal further comprises:
 a controlling unit, configured to control at least one sink terminal to play at least one to-be-played sound channel signal different from the to-be-played signal corresponding to the source terminal in the at least two to-be-played sound channel signals, to control the at least one sink terminal to cooperatively play the at least two to-be-played sound channel signals with the terminal.
10. The terminal according to claim 8 or 9, wherein the generation unit is configured to:
 obtain a movement speed of the virtual sound source and moment information of the audio signal frame; and
 determine, based on a preset movement track of the virtual sound source, the movement speed, and the moment information, the current location of the virtual sound source on the movement track.
11. The terminal according to claim 10, wherein the audio signal frame comprises music data, and the generation unit is configured to:
 determine rhythm information of music indicated by the audio signal frame; and
 determine the movement speed based on the rhythm information, wherein a faster rhythm indicated by the rhythm information indicates a faster movement speed.
12. The terminal according to claim 11, wherein the generation unit is configured to:
 determine the rhyme information based on the audio signal frame and N signal frames before the audio
- signal frame in the audio file, wherein N is an integer greater than 0.
13. The terminal according to any one of claims 10 to 12, wherein the movement track is a circle that rotates around a preset human ear location.
14. The terminal according to claim 13, wherein the terminal is the source terminal, and the source terminal or the at least one sink terminal controlled by the source terminal is located in a plane in which the circle is located.

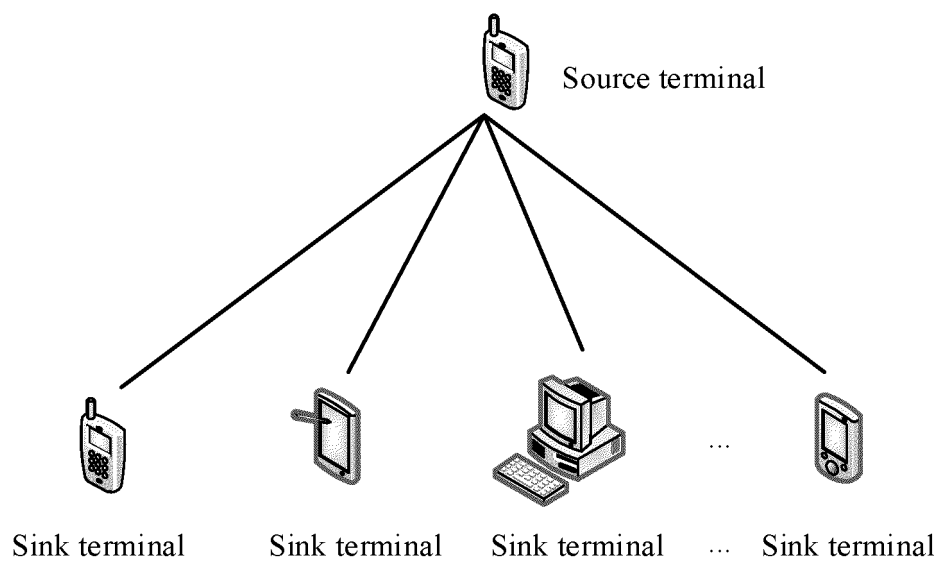


FIG. 1

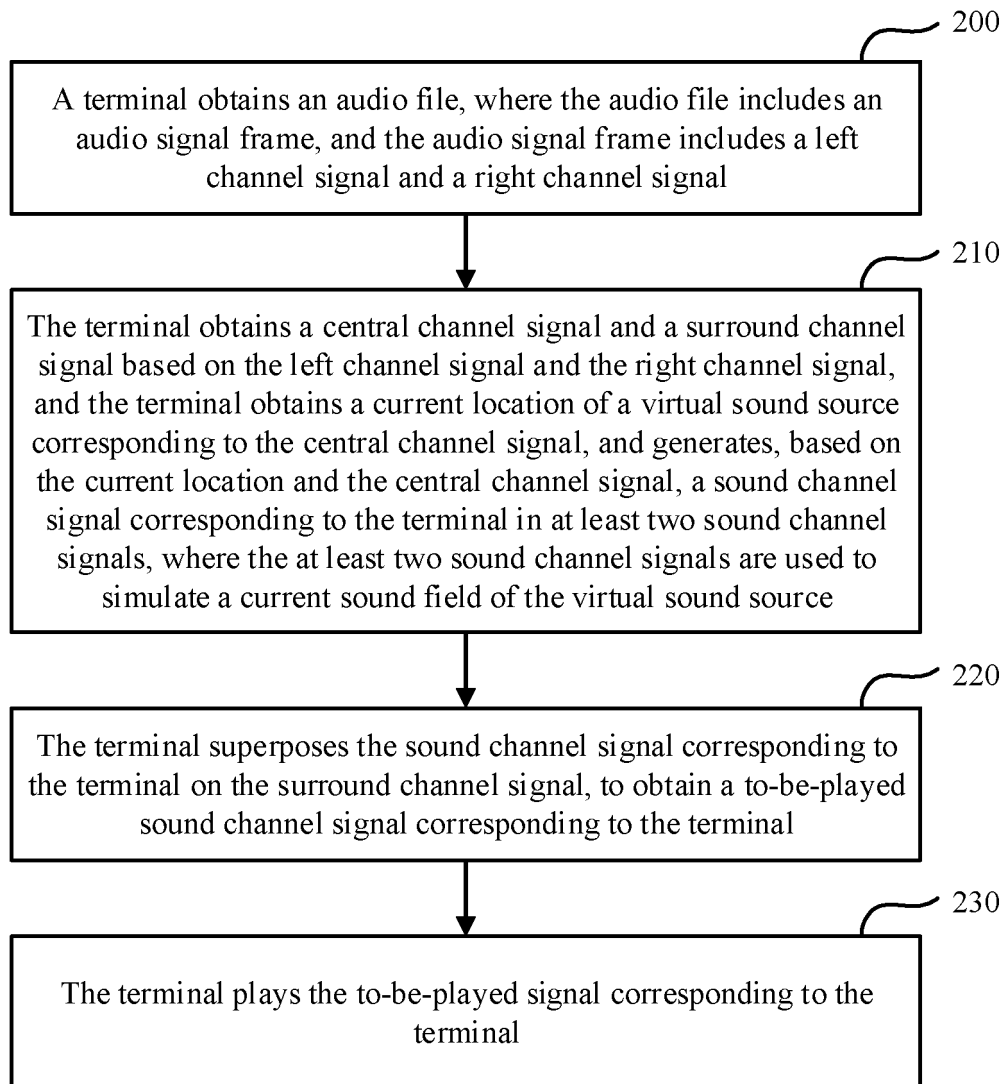


FIG. 2

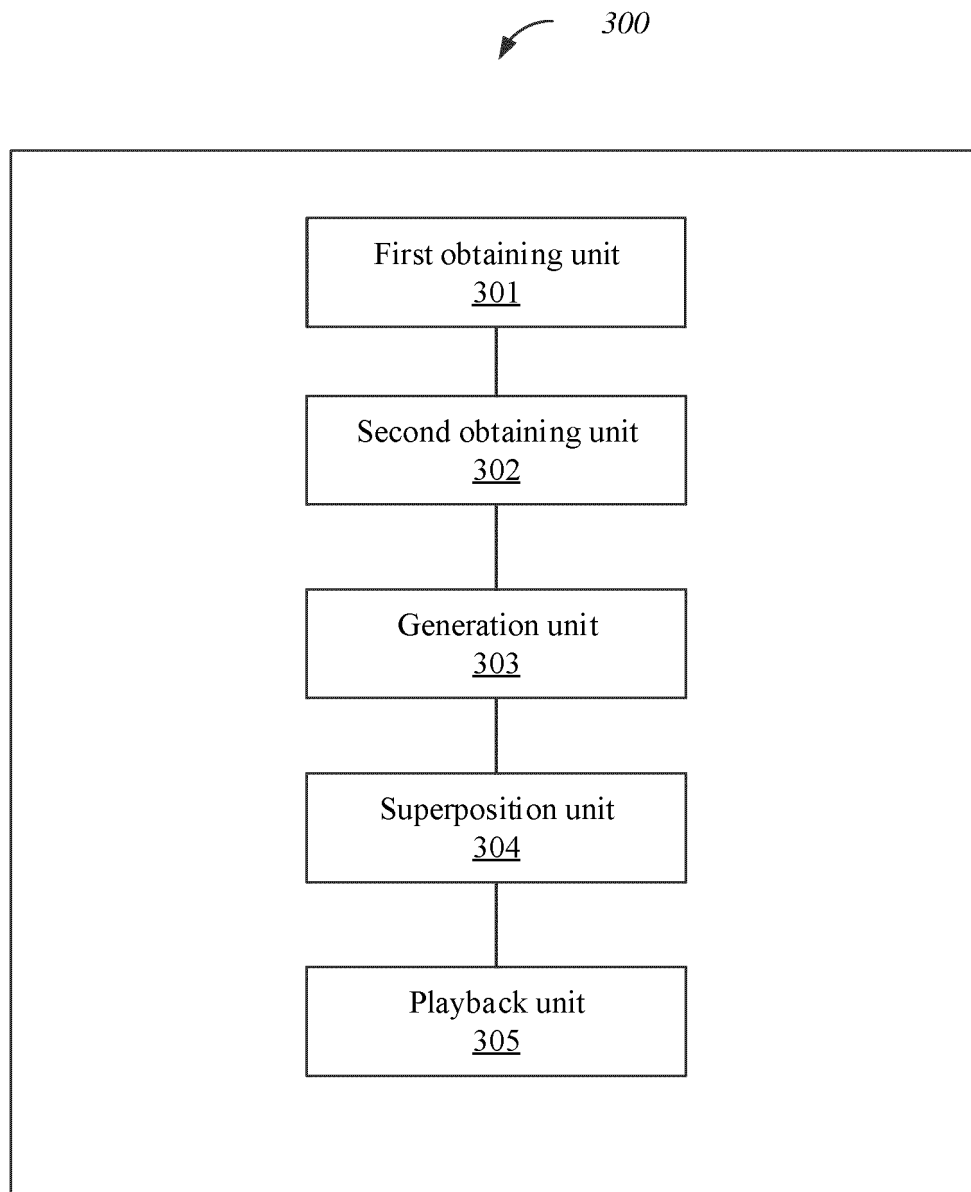


FIG. 3

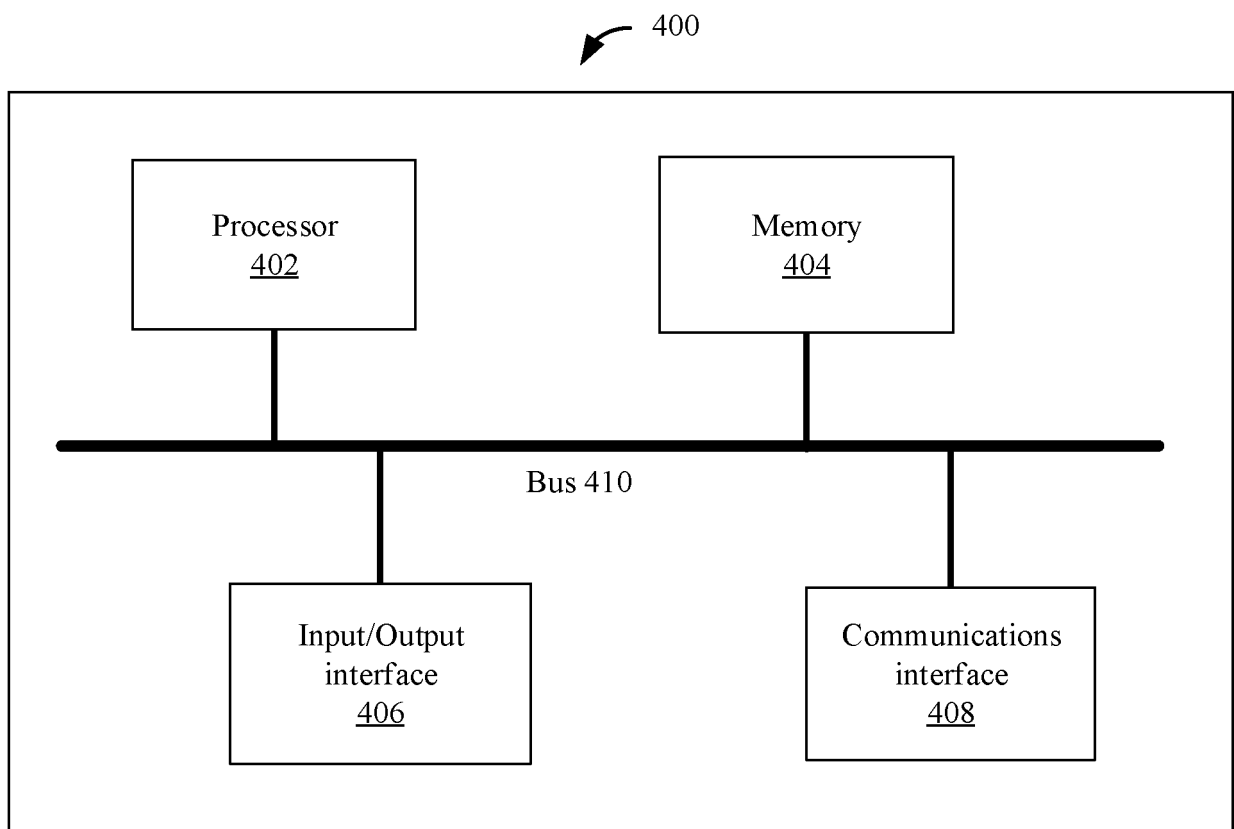


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/124244

A. CLASSIFICATION OF SUBJECT MATTER

H04R 3/00(2006.01)i; H04S 7/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04R; H04S

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN; 协同, 协作, 播放, 音频, 声道, 中置, 虚拟, 音源, 声源, 运动, cooperate, play, audio, channel, center, dummy, virtual, source, move

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106502620 A (YULONG COMPUTER TELECOMMUNICATION SCIENTIFIC (SHENZHEN) CO., LTD.) 15 March 2017 (2017-03-15) description, paragraphs [0044]-[0051], [0071] and [0088], and figures 1 and 2	1-14
Y	CN 106993249 A (SHENZHEN SKYWORTH-RGB ELECTRONICS CO., LTD.) 28 July 2017 (2017-07-28) description, paragraphs [0076]-[0087], and figures 2 and 4	1-14
A	WO 2015120633 A1 (RISHANG OPTOELECTRONICS CO., LTD.) 20 August 2015 (2015-08-20) entire document	1-14
A	WO 2015077713 A1 (VOYETRA TURTLE BEACH, INC.) 28 May 2015 (2015-05-28) entire document	1-14

☐ Further documents are listed in the continuation of Box C.
☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

22 February 2019

Date of mailing of the international search report

25 March 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2018/124244

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106502620	A	15 March 2017	None			
CN	106993249	A	28 July 2017	WO	2018196469	A1	01 November 2018
WO	2015120633	A1	20 August 2015	CN	203872353	U	08 October 2014
WO	2015077713	A1	28 May 2015	US	10015617	B2	03 July 2018
				US	2018317038	A1	01 November 2018
				US	2015181361	A1	25 June 2015
				US	2017245090	A1	24 August 2017
				US	9591426	B2	07 March 2017

Form PCT/ISA/210 (patent family annex) (January 2015)

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 201711494923 [0001]